DOCUMENT RESUME

ED 030 858 AL 001 894

By-Yeni-Komshian, Grace; And Others

A Pilot Study on the Ability of Young Children and Adults to Identify and Reproduce Novel Speech Sounds.

Johns Hopkins Univ., Baltimore, Md. Neurocommunications Lab.
Spons Agency-National Inst. of Child Health and Human Development, Bethesda, Md.
Pub Date 68
Note-17p.; Paper published in the 1968 Annual Report, Neurocommunications Laboratory, Dept. of Psychiatry, Johns Hopkins University, Baltimore, Md. 21218.
EDRS Price MF-$0.25 HC-$0.95
Descriptors-*Acoustic Phonetics, *Age Differences, Arabic, Articulation (Speech), *Auditory Perception, *Phonology, *Pronunciation Instruction, Second Language Learning

This study was designed to compare children and adults on their initial ability to identify and reproduce novel speech sounds and to evaluate their performance after receiving several training sessions in producing these sounds. The novel speech sounds used were two voiceless fricatives which are consonant phonemes in Arabic but which are completely foreign to monolingual English-speaking Americans. The sounds were presented in the context of one- and two- syllable Arabic words (target words). There were four target words, and each word was associated with a small plastic form of a common object. This approach simulated some elements of actual language learning. The objects were used to elicit identification responses from the subjects. This paper reports results for two male subjects who participated in seven half-hour training sessions given within a period of twelve days. One subject was five years old; the other was 21. Reproduction responses were judged by two adult listeners who are native speakers of Lebanese Arabic. Spectral analyses and amplitude-frequency cross-sections of the fricative portion were also made. Results do not provide any evidence that children are better than adults in acquiring novel speech sounds. (DO)

ED030858

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

A PILOT STUDY ON THE ABILITY OF YOUNG CHILDREN AND ADULTS TO IDENTIFY AND REPRODUCE NOVEL SPEECH SOUNDS

Grace Yeni-Komshian

David A. Zubin

Evangelos Afendras

AL 001 894

The preparation of this paper was supported by Contract PH 43-65-637 from the National Institute of Child Health and Human Development, National Institutes of Health.

A PILOT STUDY ON THE ABILITY OF YOUNG CHILDREN AND ADULTS TO IDENTIFY AND REPRODUCE NOVEL SPEECH SOUNDS

Introduction

The aim of this study is to examine some aspects of the commonly held view that young children are better able to learn the phonology of a second language than adults. A young child's apparent facility for acquiring new speech sounds is indirectly supported by the observation that foreign accents in speaking a second language usually begin to appear at about 11-14 years, and learning a second language becomes increasingly difficult beyond the age of 15 (1).

A learner's speech perception and production capabilities play a major role in determining the facility with which new speech sounds are acquired. This study is designed to compare children and adults on their initial ability to identify and reproduce novel speech sounds and to evaluate their performance after receiving several training sessions in producing these sounds.

Method

The novel speech sounds used in this study are two voiceless fricatives which are consonant phonemes in Arabic:

/x/: a voiceless velar fricative which includes velar "scrape" or uvular trill in allophonic variation.

/ḥ/: a voiceless pharyngeal fricative which requires strong friction without "scrape."

These speech sounds are completely foreign to the phonological system of monolingual English speaking Americans. Children, whose native tongue is Arabic, can produce them with ease. Figure 1 presents amplitude-frequency cross-sections (narrow-band, 100 cycles band width) of these sounds produced

by two female speakers of Lebanese Arabic: an adult and a five years, eight month old child (5/8). The acoustic patterns for /x/ and /ḥ/ in initial and medial position were quite similar. The cross-section displays in Figure 1 are averaged over several productions of the adult speaker, (a) and (b), and based on single productions, (c) and (d), for the child. The salient distinctions between the acoustic characteristics of /x/ and /ḥ/ produced by the adult speaker are: (1) The peak energy for /x/ is between 4.0 and 6.5 KHz, and for /ḥ/ between 2.0 and 4.0 KHz; (2) Zero energy for /x/ occurs at the peak energy frequency for /ḥ/. The displays for the Lebanese speaker (age, 5/8) indicate energy peaks at higher frequencies than the adult; however, the pattern of acoustic distinctions is the same.

The novel speech sounds were presented in the context of one-and two-syllable Arabic words, referred to as target words. There were four target words, and each word was associated with a small plastic form of a common object. This approach simulated some elements of actual language learning. The objects were used to elicit identification responses from the subjects.

The target words, transcribed phonemically, and the objects associated with each word are listed below:

Target words:	xat	daxal	ḥat	daḥal
Objects:	house	bear	girl	dog

Thus the target words were composed of two sets of minimally contrasting pairs of words in which the novel speech sounds were in initial and medial positions.

The subjects were tested individually in a sound-treated room. A loudspeaker was used to present the tape-recorded models of the target words. The

tape contained seven recordings of each target word which were taped in three sequential repetitions of each word, followed by two additional rounds of two repetitions of each word. These target words were produced by a native speaker of Lebanese Arabic who also served as the experimenter during the training sessions. The subjects were told that they were going to be taught how to say house, bear, girl and dog in a foreign language. The experimenter pointed to the appropriate objects as each target word was presented to the subject. The subject was instructed to reproduce the target words as accurately as possible. Following the reproduction task, the subject was also asked to identify the object associated with each target word. The identification test was based on tape-recorded samples of the target words and occasionally the live productions by the experimenter. The subject's task was to point to the object which he associated with the target word presented to him.

This paper reports results for two male subjects who participated in seven half-hour training sessions given within a period of twelve days. Subject MF was five years, two months, and subject RM was twenty-one years old.

During the first training session the experimenter did not attempt to shape the subjects' responses in any way. For the remaining sessions, differing amounts of live training with feedback were provided to the subjects, in addition to the reproduction and identification tests described above. The live training consisted of having a relatively rapid exchange between the experimenter, who produced the target words, and the subject, who imitated her productions. The experimenter informed the subject if his responses were getting better or not. Thus, the subject was able to observe the gross articulatory movements

required for the production of the target sounds. However, he was not given any information about specific articulatory movements.

The results described below are based on the subjects' reproduction and identification responses during the first and seventh training sessions.

Results and Discussion

The reproduction responses were judged by two adult listeners who are native speakers of Lebanese Arabic. Amplitude-frequency cross-sections were made of a selected number of reproductions from session one and seven. The results based on listener's judgments and identification responses will be presented first.

The subjects' reproductions of the taped target words during sessions one and seven were edited onto another tape which was presented to the judges for evaluation. The listening tape consisted of seven reproductions of each target word for each session. It did not include the target words that the subject was attempting to reproduce. The judges were asked to evaluate the initial sound in one-syllable responses and the medial sound in two-syllable responses. Four major classes of evaluations were obtained from the listeners' judgment: (1) an acceptable production of the novel speech sounds, (2) an approximation of /x/ or /ḥ/, not acceptable however, to native speakers, (3) substitution of [h] for both novel speech sounds, and the substitution of [x] for /ḥ/.

Table 1 presents the results based on the listeners' judgments. During session one, both subjects substituted [h] for the novel speech sounds. The reproductions of subject MF (age, 5/6) were uniformally judged as [h] by both listeners. Most of subject RM's reproductions (age, 21/0) were also judged

as [h]; however, six out of the seven reproductions of daxal were rated as partial approximations of /x/. These approximations were a modified [h], symbolized as [h-x] in Table 1. Two of subject RM's reproductions of ḥat were also rated as partial approximations, symbolized as [h-ḥ] in Table 1. Thus the older subject appeared to be better able than the younger subject to reproduce some features of the novel speech sounds during the first training session. The substitution of [h], suggests that the English phoneme /h/ bears the most similarity to the novel speech sounds.

The identification responses during the first session indicated that both subjects were unable to differentiate between xat and ḥat, and between daxal and daḥal. However, they were able to identify those objects associated with the one-syllable target words and those associated with the two-syllable target words.

In session seven, the subjects' reproductions of the target words were markedly different. They reproduced the novel speech sounds in xat and daxal in a manner that was acceptable to the judges. However, both had difficulty with /ḥ/. Most of subject MF's reproductions were rated as approximations of /ḥ/. The remaining reproductions were rated as [h]. The adult subject, RM, had one reproduction of ḥat which was judged as acceptable by the listeners. Although this subject did not produce [h] during the seventh session, most of his reproductions of target words with /ḥ/ had velar "scrape" which were rated as unacceptable approximations of /x/, or, acceptable /x/ by the listeners.

The subjects' identification of the target words improved greatly. During the seventh session they were presented with a 40-item identification test tape.

The first 10 items contrasted xat with ḥat, the second set of 10 items contrasted daxal with daḥal, and the remaining items had all four target words in random order. Subject MF had seven errors, six of which occurred on the last 20 items. The adult subject, RM, had 2 errors which also occurred on the last 20 items.

These results do not provide any evidence indicating that children are better than adults in acquiring novel speech sounds. The novel speech sound /x/ was easier to learn than /ḥ/. The subjects were able to produce xat and daxal in a consistently acceptable way. They appeared to have different renditions of ḥat and daḥal, suggesting that even after seven training sessions the subjects did not have a stable awareness of the critical articulatory movements necessary for the production of /ḥ/. The fact that subject MF did not substitute [x] for the fricatives in ḥat and daḥal might reflect an important articulatory feature that is found in children. The adult subject was aware of substituting [x] for /ḥ/. He succeeded in producing one acceptable [ḥ], but he was unable to repeat the necessary articulatory gestures.

The subjects' reproduction responses were analyzed spectrographically (wide-band, 400 cycles band width) and amplitude-frequency cross-sections of the fricative portion were made. The cross-section display represents an average over 20 msec. The best reproductions of each of the four target words in session seven and their corresponding matches in session one were thus analyzed. The results are presented in Figures 2, 3, 4 and 5. In each figure the target words and the listeners' ratings of the reproductions are listed below the tracings of the amplitude-frequency displays. The frequency range is zero to 7.0 KHz, and the amplitude scale is relative db (1mm=1db).

Figures 2 and 3 present the cross-sections for subject MF's responses in session one and seven respectively. The four sections in Figure 2 are quite similar, all having peak energies at about 1.5 KHz. The sections in Figure 3 illustrate the magnitude of change in subject MF's responses. The peak energy for responses rated as acceptable [x], sections (a) and (b), are characterized by energy peaks at about 0.5 and 2.0 KHz, with relatively less energy at about 1.0 KHz. Section displays (c) and (d) represent two portions of MF's reproduction of hat. He produced a prolonged fricative the first portion of which, section (c), was an acceptable approximation of /ḥ/ and the last portion, section (d), was rated as [h]. Section (e) is quite similar to section (c) in th frequency range 0.5 to 3.0 KHz; however, in the frequency range above 3.0 KHz, section (e) shows a wider spread of energy. This might explain why section (e) was rated as a partial approximation of /ḥ/ while the portion in section (c) was acceptable to the listeners.

The acoustic analysis of subject MF's responses clearly illustrate the change in his reproductions. On the seventh training session he was able to produce two novel speech sounds, ([x] and [h-ḥ]), one of which was acceptable to the native listeners.

Figures 4 and 5 present cross-section displays for subject RM's responses in session one and seven respectively. The two cross-sections rated as [h], Figure 4 (a) and (d), have peak energies at about 1.0 KHz and no energy above 2.0 KHz. Sections (b) and (c) contain some energy between 3.0 and 4.0 KHz. They were rated as partial approximations of /x/ and /ḥ/ respectively. The cross-sections of RM's responses in session seven, Figure 5, are

characterized by a wider spectrum of energy. Sections (a) and (b) in Figure 5 display the patterns for responses rated as acceptable [x]. Energy peaks at about 1.5, 4.5 and 6.0 KHz are present in both sections. In addition, a reduction or absence of energy is noted at about 2.0 KHz. Section (c) in Figure 5 was rated as an acceptable production of /ḥ/, while section (d) was not acceptable because of velar "scrape." The main difference between these two sections was the presence of energy at higher frequency ranges in section (d).

The results of the cross-section analysis demonstrated the change in subject RM's responses. This subject, like MF, was also able to produce two novel speech sounds, one of which was an acceptable rendition of the Arabic phoneme /x/.

These results support the ratings obtained from the listeners' judgments. Although the spectral displays show two different patterns which correspond with each of the novel speech sounds produced by each subject, they do not correspond across the two subjects.

A more detailed examination of the acoustic features of the novel speech sounds is necessary to fully evaluate the responses. Such information would provide an objective measure of the subject's progress in learning novel speech sounds. A detailed analysis of the acoustic features of /x/, /ḥ/ and /h/ is in progress.

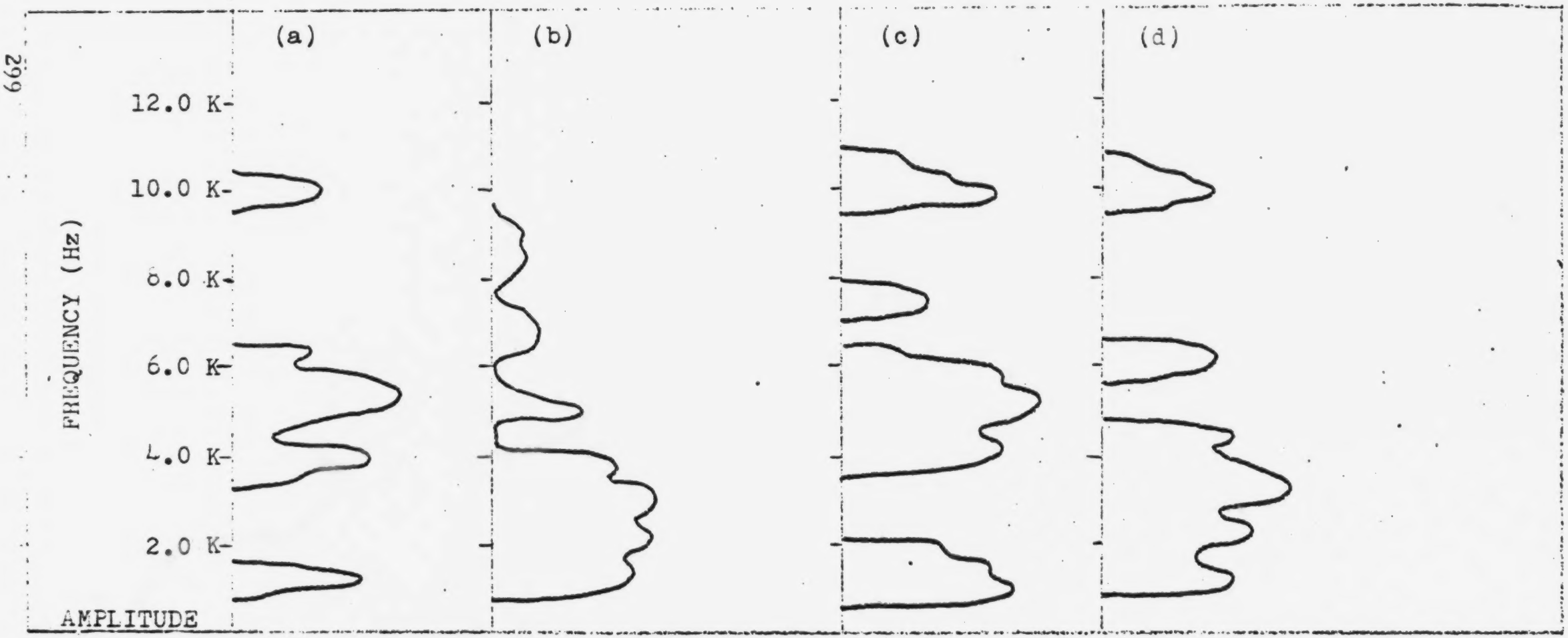


Figure 1. Amplitude-frequency cross-sections of /x/ and /ḥ/ produced by two female speakers of Lebanese Arabic. The speech samples were analyzed at half-speed.

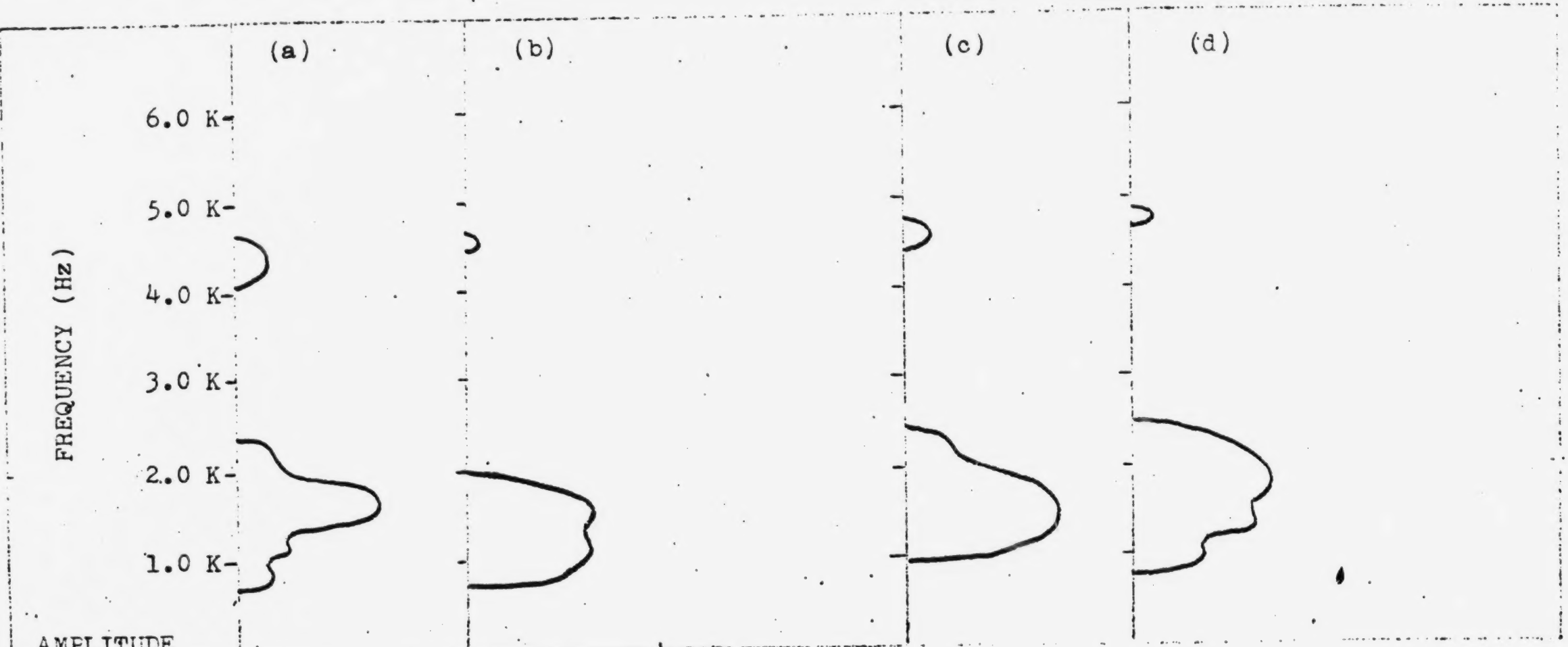
Table 1

Summary of the Listeners' Judgements

	Target Words	Session I Reproduction	n	Session 7 Reproduction	n
Subject MF	xat	[h]	7	[x]	7
	daxal	[h]	7	[x]	7
Age 5/2	ḥat	[h]	7	[h]	2
				[h-ḥ]	5
	daḥal	[h]	7	[h]	2
				[h-ḥ]	5
Subject RM	xat	[h]	7	[x]	7
	daxal	[h]	1	[x]	7
		[h-x]	6		
Age 21/0	ḥat	[h]	5	[ḥ]	1
		[h-ḥ]	2	[h-x]	2
				[x]	4
	daḥal	[h]	7	[h-x]	4
				[x]	3

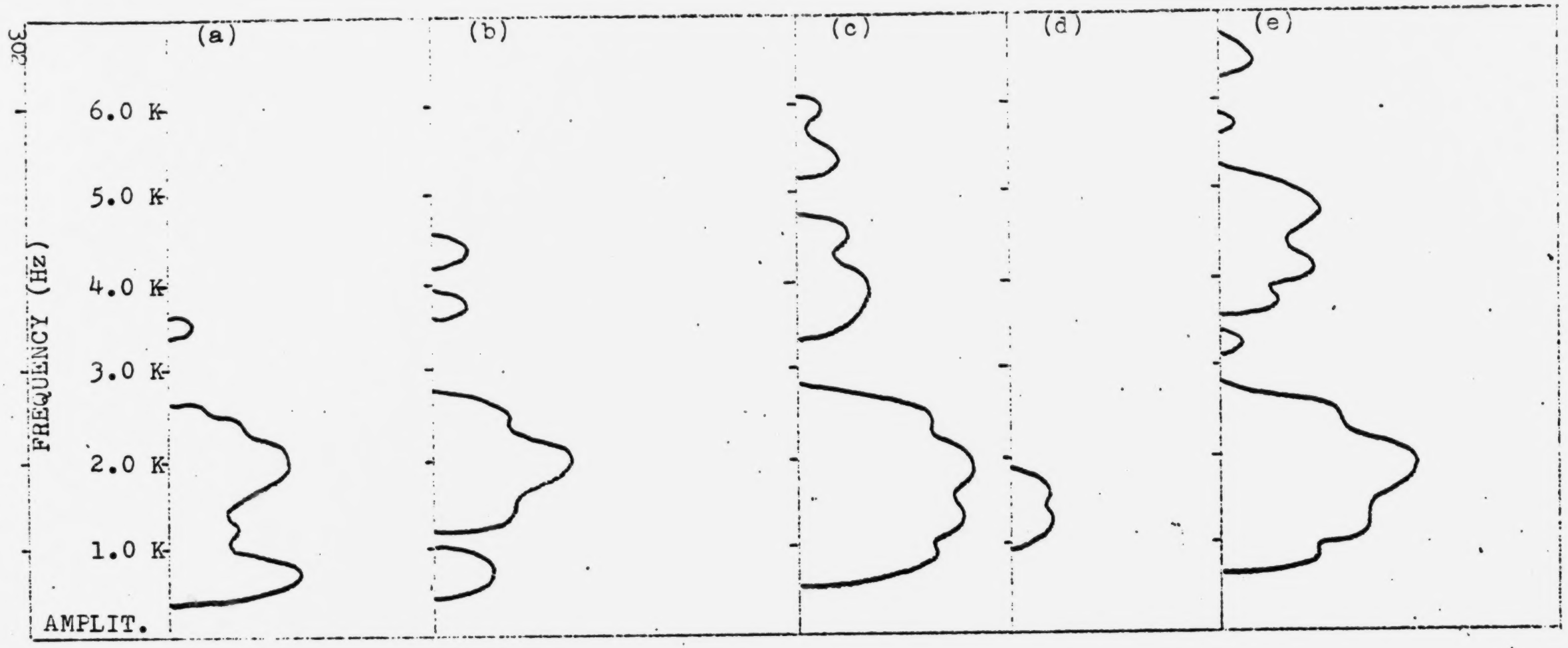
[h-ḥ] = unacceptable approximations of [ḥ] .

[h-x] = unacceptable approximations of [x].



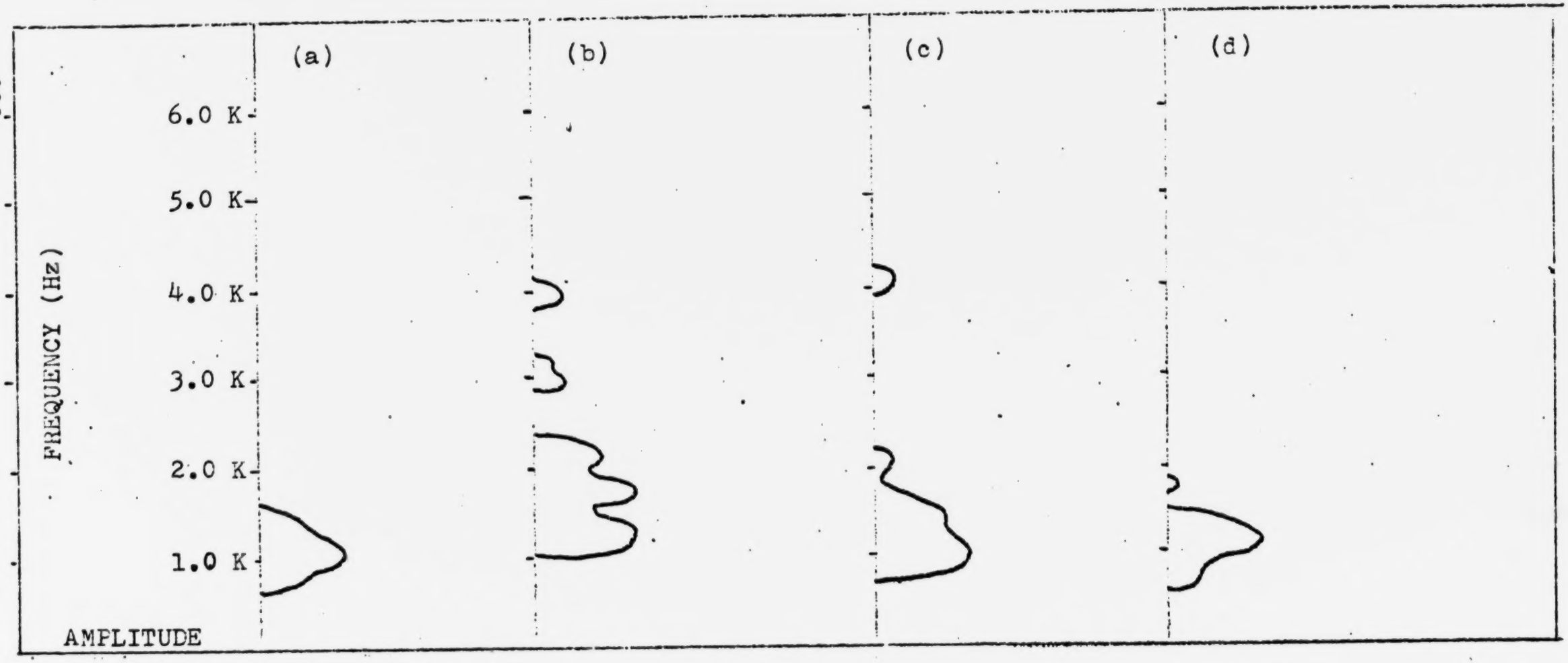
Target word:	xaṭ	daxal	ḥat	daḥal
Reproduction rated as :	[h]	[h]	[h]	[h]

FIGURE 2. Amplitude-frequency cross-sections for the responses of subject MF, age 5/2, during training session one.



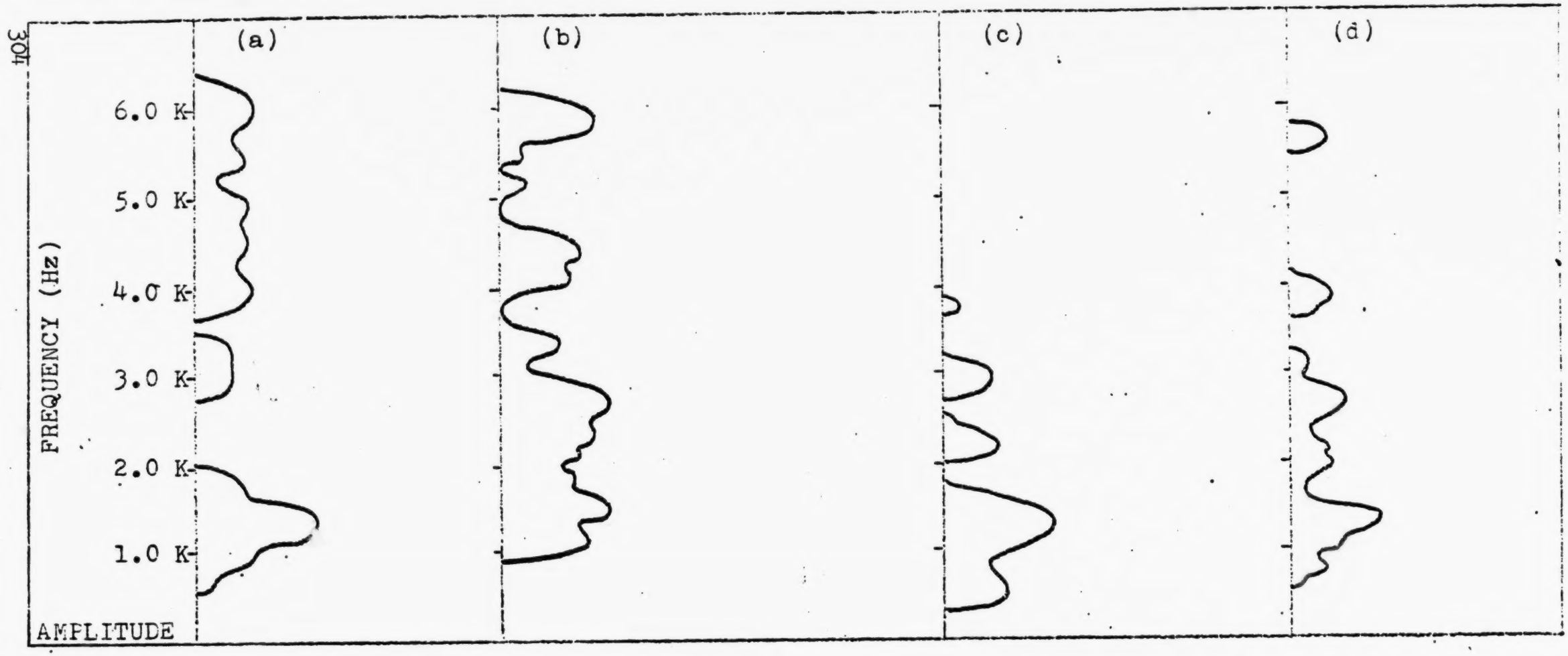
Target word:	xat	daxal	ḥat	ḥat	daḥal
Reproduction rated as :	[x]	[x]	[h——h]		[h-ḥ]

FIGURE 3. Amplitude-frequency cross-sections for the responses of subject MF, age 5/2, during training session seven.



Target word:	xat	daxal	ḥat	daḥal
Reproduction rated as :	[h]	[h-x]	[h-ḥ]	[h]

FIGURE 4. Amplitude-frequency cross-sections for the responses of subject RM, age 20/0, during training session one.



Target word:	xat	daxal	ḥat	daḥal
Reproduction rated as :	[x]	[x]	[ḥ]	[h-x]

FIGURE 5. Amplitude-frequency cross-sections for the responses of subject RM, age 21/0, during training session seven.

REFERENCES

1. Lenneberg, E.H. Biological Foundations of Language. New York: John Wiley & Sons, 1967.